

Development and applications of the FV3 GEOS-5 adjoint modeling system

Daniel Holdaway (dan.holdaway@nasa.gov)

NASA Global Modeling and Assimilation Office
Goddard Earth Sciences, Technology and Research

with

Jong G. Kim, Shian-Jiann Lin, Ron Errico, Ron Gelaro, James Kent,
Larry Coy, Jim Doyle, Alex Goldstein

NRL Seminar
Wednesday 17 May, 2017



Introduction

Over the last few years GMAO has been developing the tangent linear and adjoint of GEOS-5.

This system supports a number of current and upcoming operational and research capabilities:

- Daily calculation of observation impacts (FSOI)
- 4DVAR data assimilation using GSI
- Adjoint sensitivity to initial conditions
- Singular vector calculations
- Assessment of new observations and changes to current
- Estimation of surface fluxes of emissions
- Analyze ensemble techniques

Introduction

We have a nonlinear model for predicting the future atmosphere,

$$y_i = m(x_1, x_2, \dots, x_j),$$

where x and y are discrete model variables at an old and new time.

From a predictability perspective we may be interested in:

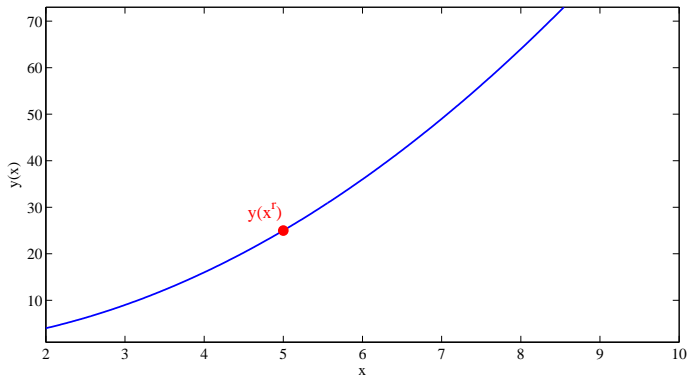
1. How does a forecast change with respect to the inputs? If there's some error or change due to an observation, how does that quantity grow?
2. What caused some aspect of the forecast we're analyzing, e.g. an error? What specifically was it about the initial state that led to this? What was the value of an observation to the forecast?

Introduction

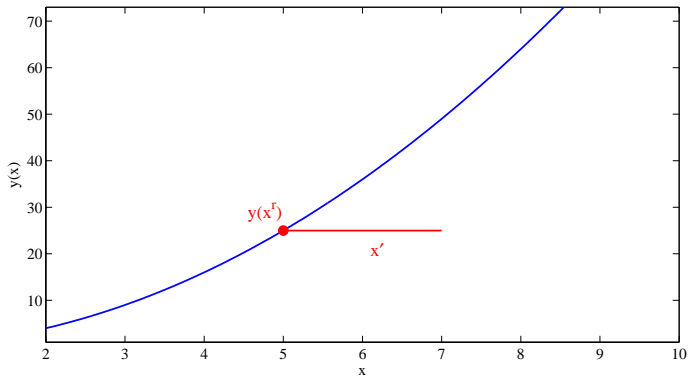
We're trying to get a handle on perturbations so we 'linearize' model variables by separating into reference and perturbation parts,

$$x_j = x_j^r + x_j'.$$

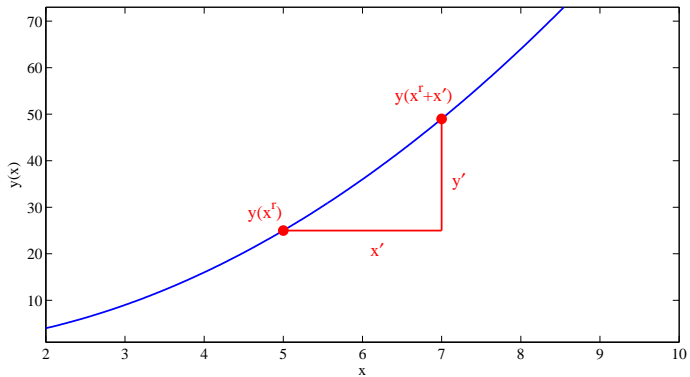
Introduction



Introduction



Introduction



Introduction

We're trying to get a handle on perturbations so we 'linearize' model variables by separating into reference and perturbation parts,

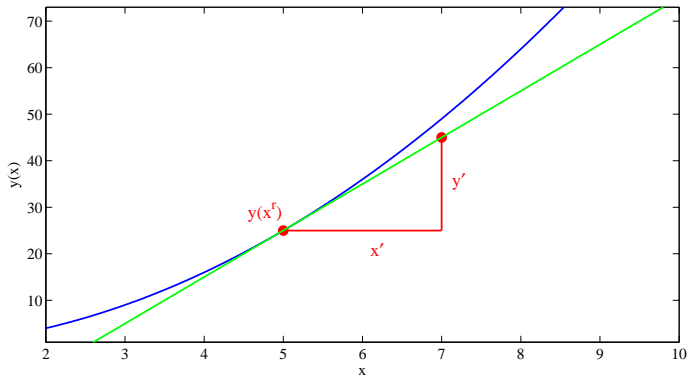
$$x_j = x_j^r + x_j'.$$

The output perturbation can be approximated by Taylor expansion,

$$\begin{aligned} y_i' &= m(x_1^r + x_1', x_2^r + x_2', \dots) - m(x_1^r, x_2^r, \dots), \\ &= \cancel{m(x_1^r, x_2^r, \dots)} + \frac{\partial y_i^r}{\partial x_1} x_1' + \frac{\partial y_i^r}{\partial x_2} x_2' + \dots - \cancel{m(x_1^r, x_2^r, \dots)}, \\ y_i' &= \sum_j \left(\frac{\partial y_i}{\partial x_j} \right)^r x_j'. \end{aligned}$$

The above equation is the **tangent linear model (TLM)** and gives an approximation for the growth of perturbations x_j' .

Introduction



Introduction

The TLM is a useful tool but doesn't help us with the second question. Introduce a scalar measure of the outputs, $J = J[\mathbf{y}(\mathbf{x})]$.

Using the Taylor series again,

$$J' = \sum_i \frac{\partial J^r}{\partial y_i} y'_i \quad \text{or} \quad J' = \sum_j \frac{\partial J^r}{\partial x_j} x'_j.$$

The two estimates are equivalent when both functions are linear. J is a function of a function so expand with chain rule,

$$\frac{\partial J^r}{\partial x_j} = \sum_i \left(\frac{\partial y_i}{\partial x_j} \right)^r \frac{\partial J^r}{\partial y_i}.$$

Obtain a powerful equation from end time to beginning time.

Introduction

The TLM sums $\partial y_i / \partial x_j$ over j while this new model sums over i . Adopting vector notation to compute all y'_i or $\partial J / \partial x_j$,

$$\mathbf{M}_{i,j} = \frac{\partial y_i}{\partial x_j}.$$

For the TLM sum over the columns, for this new model sum over rows. In standard matrix multiplication we sum over columns. The TLM is fine,

$$\mathbf{y}' = \mathbf{M}\mathbf{x}',$$

but for the new model we have to transpose (adjoint) the matrix,

$$\frac{\partial J}{\partial \mathbf{x}} = \mathbf{M}^\top \frac{\partial J}{\partial \mathbf{y}}.$$

Outline

- Development of the GEOS-5 adjoint
- Applications
- Ongoing & future work

Overview of the system

The GEOS-5 adjoint consists of:

- Finite volume cubed sphere (FV3) dynamical core
- Relaxed-Arakawa convection
- Bacmeister single moment cloud scheme
- Turbulence based on Lock et al closure
- Gravity wave drag
- Long- and short-wave radiation
- GOCART dust physics

FV3 dynamical core

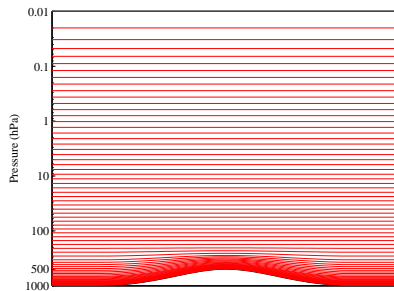
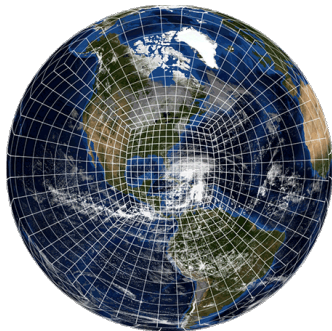
The FV3 dynamical core is developed by NOAA GFDL and has been central to multiple versions of GEOS.

- Finite volume numerics
- Cubed-sphere grid
- D-grid shallow water dynamics
- Lagrangian vertical coordinate with PPM remapping to σ - p
- PPM horizontal advection schemes (around 15 options)

Through NGGPS saw the following updates,

- Updated non-hydrostatic algorithm
- Single precision (30% faster)
- Nested grid capability

FV3 dynamical core



Strong versus weak nonlinearity

Dynamical cores largely exhibit only 'weak' nonlinearity. By this we mean that the following condition holds,

$$\lim_{\Delta \mathbf{x} \rightarrow 0} \mathbf{m}(\mathbf{x} + \Delta \mathbf{x}) - \mathbf{m}(\mathbf{x}) = \mathbf{M}\Delta \mathbf{x}$$

Physics schemes exhibit 'strong' nonlinearity. Due to the use of piecewise (linear or otherwise) functions,

$$\lim_{\Delta \mathbf{x} \rightarrow 0} \mathbf{m}(\mathbf{x} + \Delta \mathbf{x}) - \mathbf{m}(\mathbf{x}) \neq \mathbf{M}\Delta \mathbf{x}$$

In earlier NWP models it was possible to have correlations of almost 1 between tangent linear and nonlinear dynamical cores out to 5 days or so.

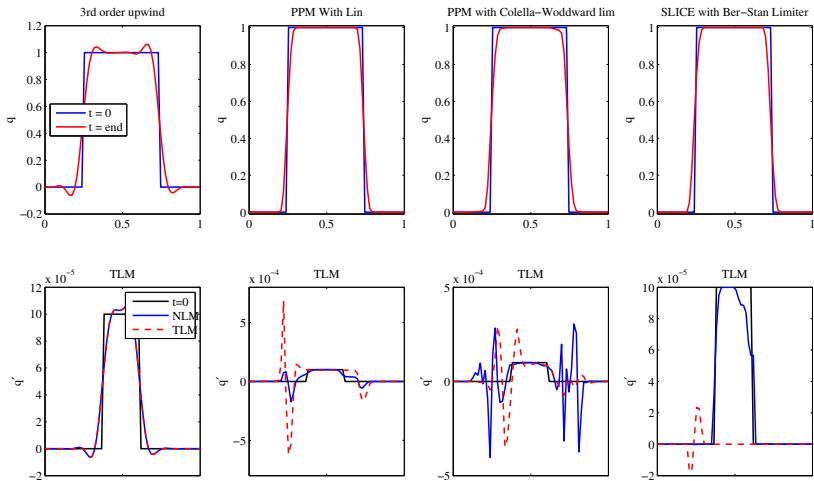
Strong nonlinearity in dynamical cores

The primitive equations of the atmosphere are weakly nonlinear.

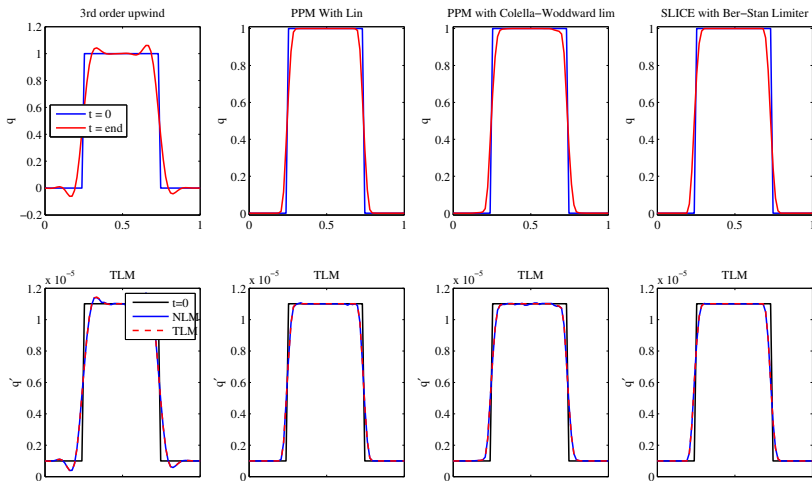
However, these equations are not solved in their continuous form but discretised and solved numerically, with a finite volume method in FV3's case. In doing so we run into problems, e.g. *linear positive definite advection schemes cannot be greater than first order* (Godunov, 1959).

As a result sophisticated nonlinear shape preserving schemes are employed to solve the horizontal and vertical (remapping) transport.

Strong nonlinearity in dynamical cores

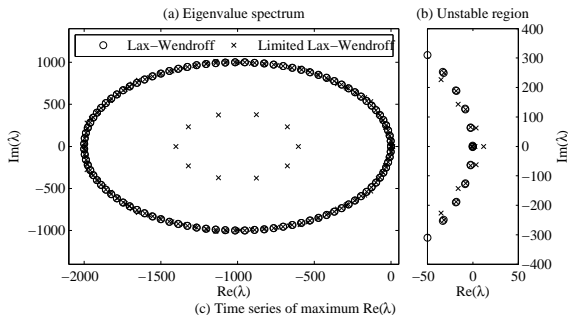


Strong nonlinearity in dynamical cores



Strong nonlinearity in dynamical cores

These strong nonlinearities generally also produce linear instability and thus uncontrolled growth of the perturbations.



A 2D spherical geometry test case using prescribed winds and true solutions has been developed to test the approach to linearized advection in NWP (Kent and Holdaway, 2017, QJRM).

Transport in the FV3 adjoint

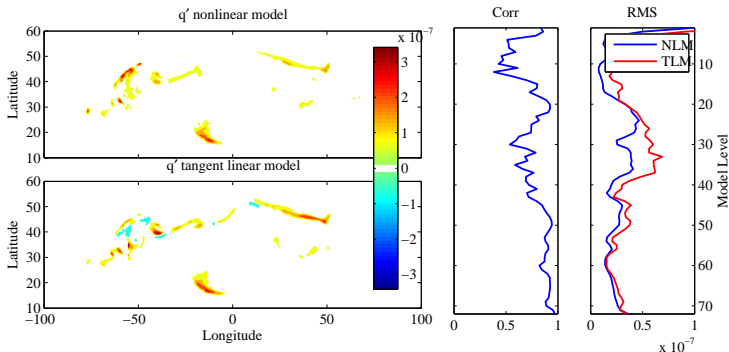
All positive definite PPM schemes available in FV3 are found to exhibit poor tangent linear behavior. Instead transport is handled by purely linear differentiation (weakly nonlinear overall due to varying winds).

A third order method is introduced to provide a scheme with more diffusion than the fifth order linear scheme present in FV3. In the sponge layer first order advection is used.

For remapping a first order linear interpolation is implemented for all perturbation quantities. Cubic was also tested but not found to produce good results.

Transport in the FV3 adjoint

5 day integration of a slotted cylinder tracer perturbation with wind perturbations set to zero. Third order linear differentiation for the tracer and linear remapping.



May be beneficial to add some targeted vertical hyper-diffusion to the tracer perturbations.

Damping in the FV3 adjoint

Divergence, vorticity and external mode damping in FV3 are tuned in concert with the transport scheme. Some schemes are more diffusive than others, requiring less damping to maintain stability.

For the linearized version of FV3 significantly more damping is required to maintain stability. Typically we need to run with second and fourth order divergence damping coefficients maxed out and with voracity damping set quite high.

Reference - perturbation splitting

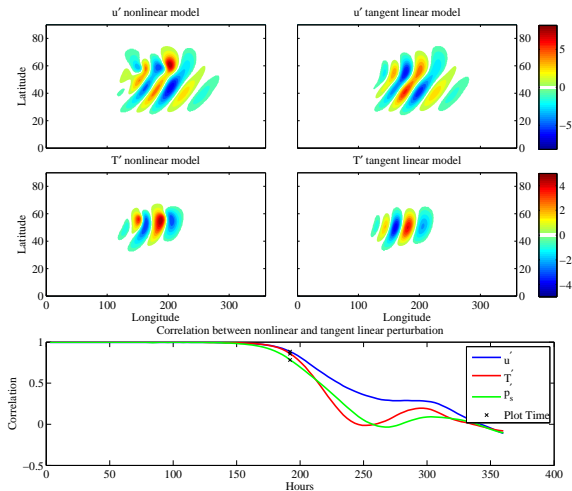
We have three choices when running the adjoint:

- Run both reference and perturbations with the transport and damping schemes that work well for the perturbations.
- Split the transport so reference values use the nonlinear schemes and perturbations use the linear schemes but both use the divergence damping that works for the perturbations.
- Split both the transport and damping schemes.

$$\begin{array}{ccccc} x' & = & M_b & \longleftarrow & M_{LIN} & \longleftarrow & M_a & x' \\ & & \swarrow & & \searrow & & & \\ & & x^r & & M_{NLIN} & & x^r & \end{array}$$

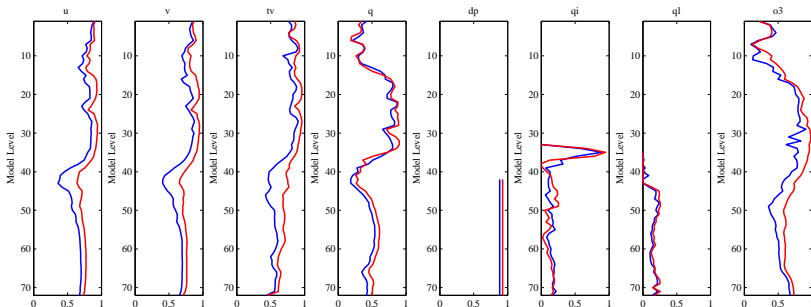
Splitting is slower but more accurate, useful for longer runs looking at sensitivity or FSOI. For 4DVAR and singular vector calculations it may be beneficial to turn splitting off.

Testing: Jablonowski-Williamson Baroclinic Instability



Testing: Analysis Increment Perturbation

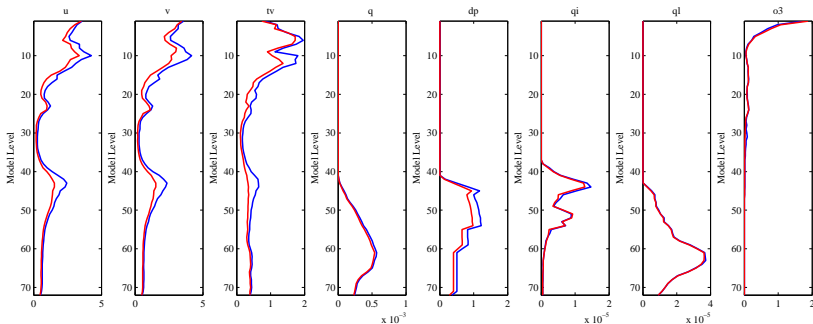
Correlation between the 24 hour nonlinear and tangent linear perturbation trajectories. Blue = old TLM, red = new TLM.



$$u' = 0.81, v' = 0.81, T_v' = 0.79, q' = 0.51, \Delta p' = 0.92.$$

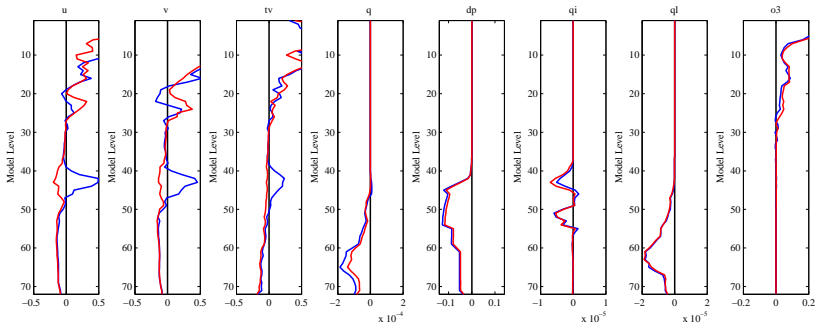
Testing: Analysis Increment Perturbation

Root mean squared error between the 24 hour nonlinear and tangent linear perturbation trajectories.



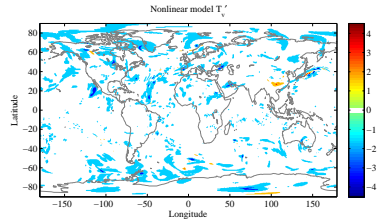
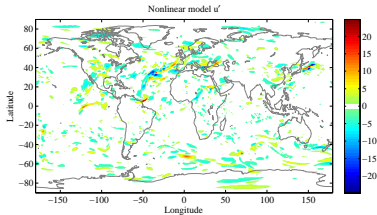
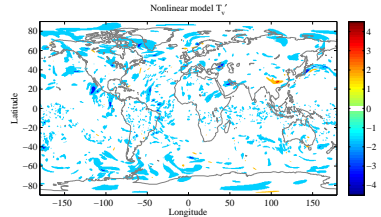
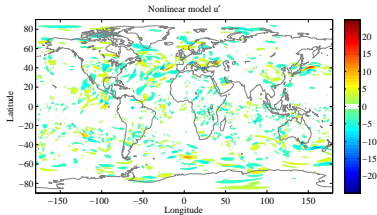
Testing: Analysis Increment Perturbation

Root mean square compared to nonlinear model.



Testing: Analysis Increment Perturbation

100hPa nonlinear and tangent linear model perturbation fields.



Autodifferentiation

The Tapenade tool is used to generate the majority of the code.

Pros:

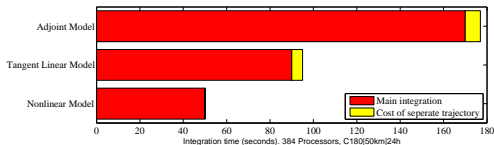
- Very flexible and works on modular & object form of FV3
- Fast, generates FV3 TLM and adjoint in matter of minutes
- Minimal re-computation (fast)
- Fast writing of checkpoints to stack using malloc()
- Inexpensive

Cons:

- Minimal re-computation (memory intensive, or how to bring down a node in one easy step!)
- Precision determined at generation time

Timing

The adjoint of FV3 is about 4 times slower than the nonlinear model, the tangent linear is around 2 times slower.



FV3 has a relatively small timestep, 75s when $\Delta x = 50\text{km}$. On the other hand it resolves small scale features very well, one the reasons it was picked in the NGGPS testing.

The only application for which this is a major concern is 4DVAR, but there are things we might be able to do.

Autodifferentiation

We have written our own push/pop checkpointing routines, allowing for compile time precision choice. This has other benefits:

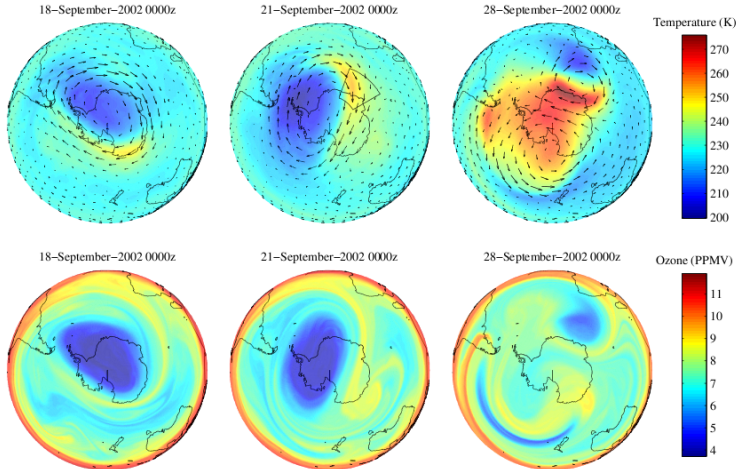
- Flexibility of memory handling, malloc/SHMEM/allocate.
- Check on superfluous saves, especially around communication.
- Ability to hold checkpoints in memory. This could save a tremendous amount of time in 4DVAR if we can spare the memory.

Outline

- Development of the GEOS-5 adjoint
- Applications
- Ongoing & future work

Sudden Stratospheric Warming

In 2002 a rare southern hemisphere sudden stratospheric warming occurred, disrupting the polar vortex and temporarily eliminating the ozone hole.

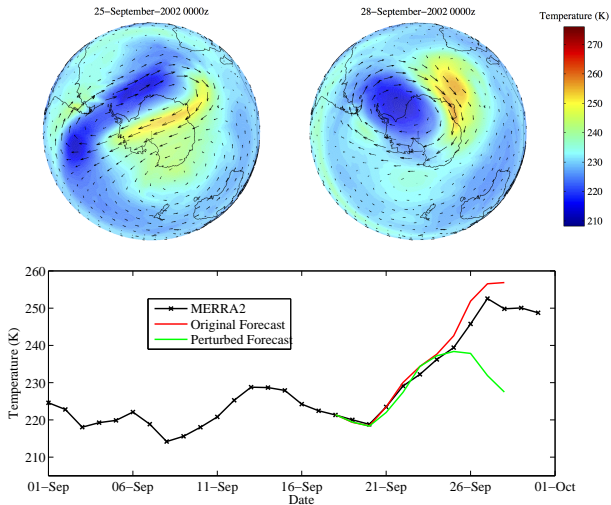


Sudden Stratospheric Warming

The adjoint model serves as a useful tool for examining the sensitivity to initial conditions ahead of the event. The method works as follows:

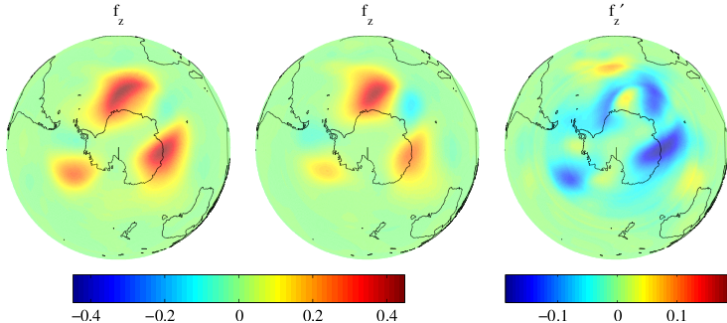
- Choose an initial metric J , here a small area at 10hPa where the warming first occurs.
- Run the adjoint backwards to find sensitivity to initial conditions 2-3 days before.
- Generate a perturbation that minimizes J using a Lagrange multiple technique.
- Run a new nonlinear forecast with initial conditions altered by the optimal perturbation.
- Verify the perturbation.
- Examine how the perturbation evolves.

Sudden Stratospheric Warming



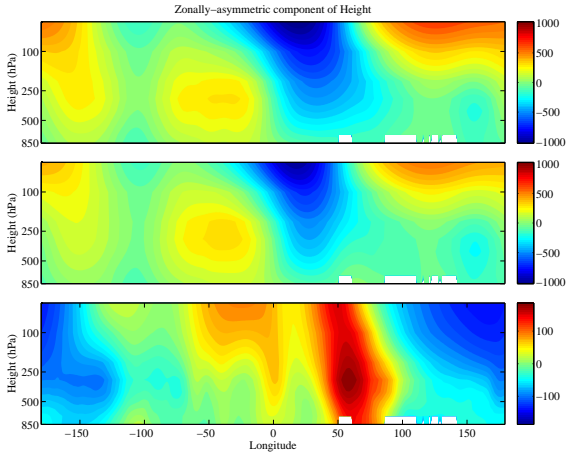
Sudden Stratospheric Warming

5 day running mean of the vertical component of the Plumb wave flux centered 24 hours before the first significant warming occurs.



Sudden Stratospheric Warming

5 day running mean of the zonally asymmetrical height centered ahead of the warming.



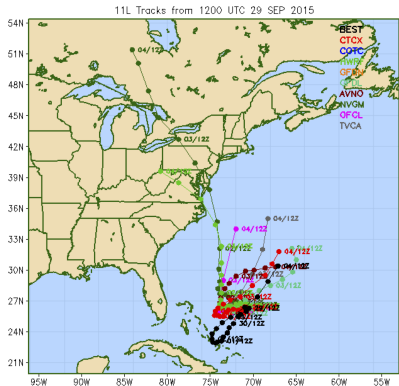
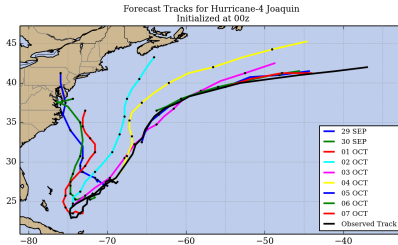
Sudden Stratospheric Warming

The sudden stratospheric warming is shown to be driven by Rossby waves emanating from tropospheric disturbances.

This is further confirmed by examining the ERA-20C reanalysis. Here only surface observations are used to derive a century long atmospheric state that includes the stratosphere. The 2002 southern hemisphere sudden stratospheric warming is evident in the record, though a little weaker than it was in reality. Further, there is evidence for 3 more southern hemisphere sudden stratospheric warmings, all occurring in the early 20th century.

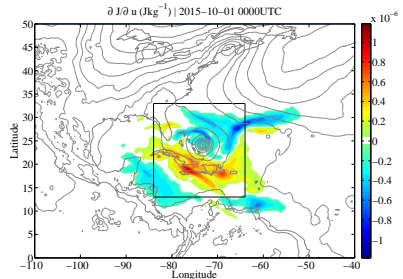
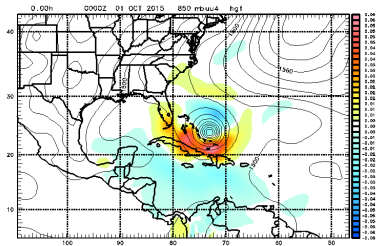
Hurricane Joaquin

Hurricane Joaquin was a category 4 Atlantic hurricane in 2015 that models incorrectly forecast as making landfall on the Eastern seaboard.

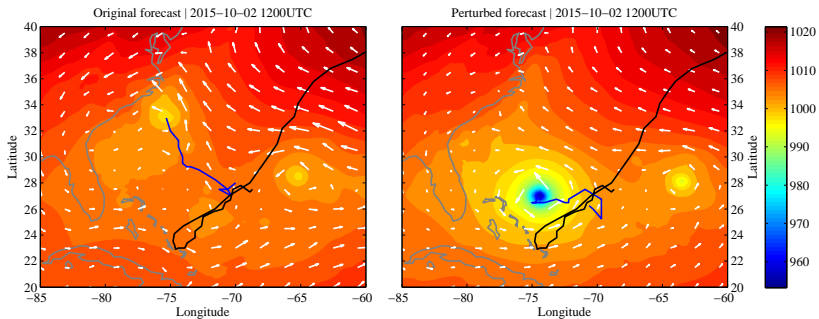


Hurricane Joaquin

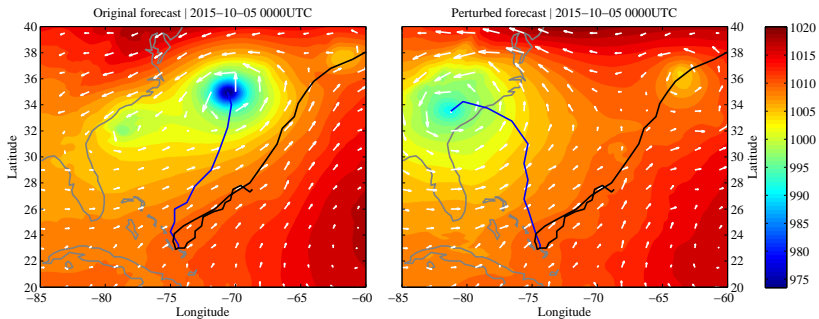
36 hour sensitivity of kinetic energy to the initial wind fields,
COAMPS versus GEOS-5 adjoint models.



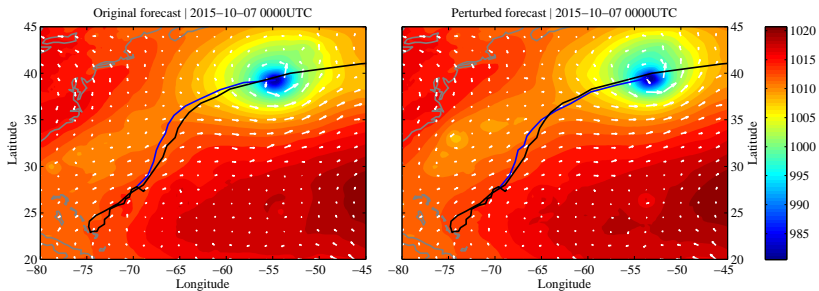
Hurricane Joaquin



Hurricane Joaquin



Hurricane Joaquin



Hurricane Joaquin

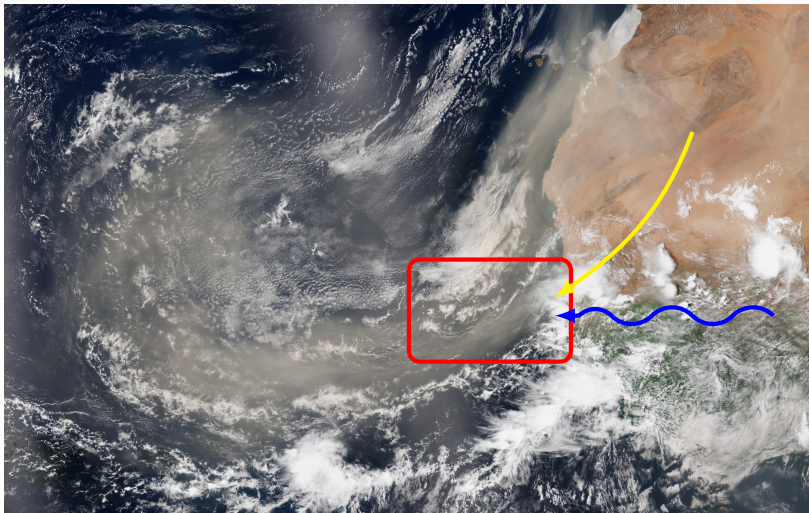
Accurate forecasting Joaquin was immensely difficult due to the high degree of sensitivity to the initial conditions.

Small changes to the initial conditions results in changes to the intensity which changes the track significantly, leading to even greater changes in intensity.

Sensitivity to Saharan dust in hurricane formation

TC formation requires heat and moisture, which can be affected by dust in a number of ways:

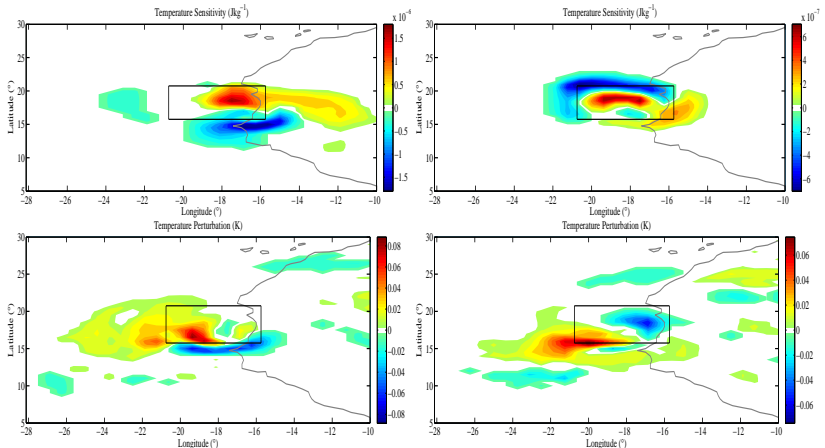
- Dust absorbs solar radiation and warms up relative to the surroundings.
- However, this prevents some solar radiation reaching the surface, cooling below the dusty layer.
- Dusty air tends to be dry and as the SAL is entrained it can reduce the available energy.
- Dust can impact condensation through micro-physical processes.



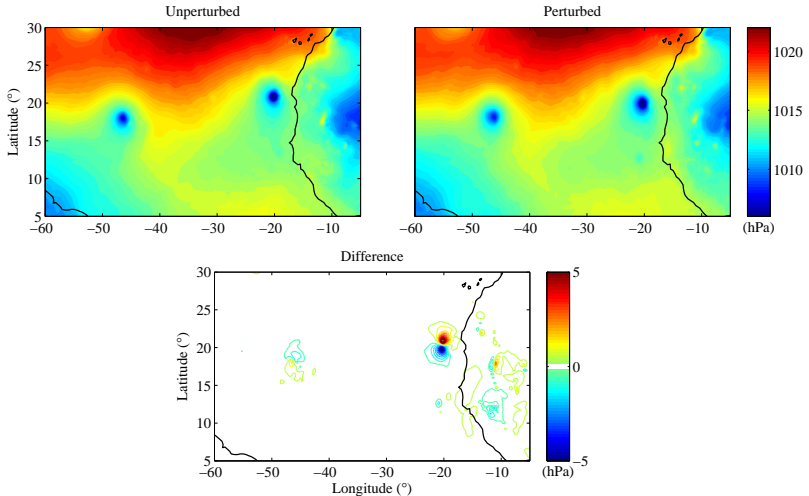
VIIRS on Suomi NPP 31-July-2013

Perturbation vs. sensitivity evolution

Evolution of temperature sensitivity evolves very differently than changes in temperature due to an optimal dust perturbation.



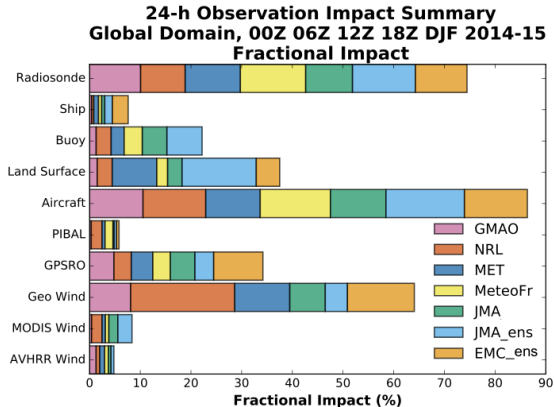
Optimal dust removal



- Helene (left) weakens and moves north.
- Other storm (right) strengthens and moves south.

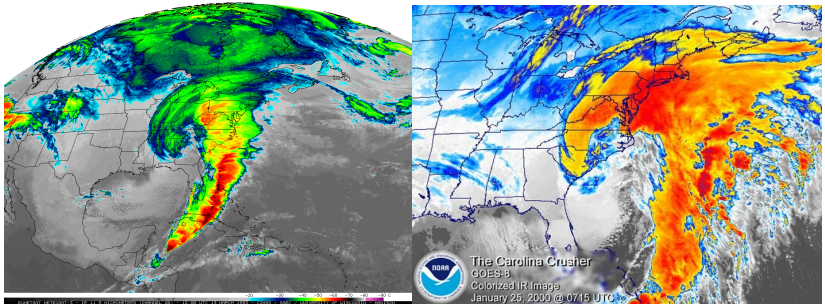
FSOI

GMAO is a partner in the international FSOI inter-comparison project with JMA, NRL, EMC, Meteo France and the Met Office. We have run two seasons of 4 times daily adjoint observation impacts.



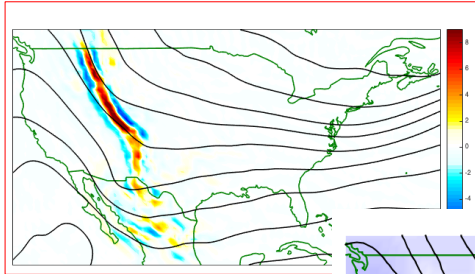
Winter storm evolution (U of Wisconsin)

Winter storms of 1993 (predictable) and 2000 (unpredictable).



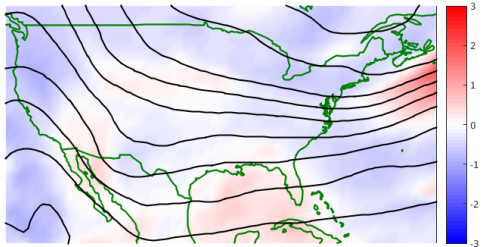
Winter storm evolution (U of Wisconsin)

1993 storm of the century case



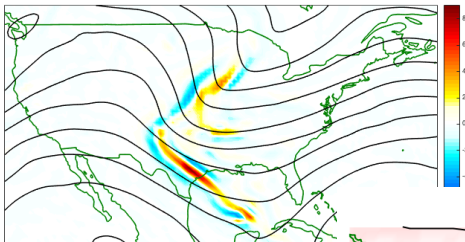
ψ_{500}

F000



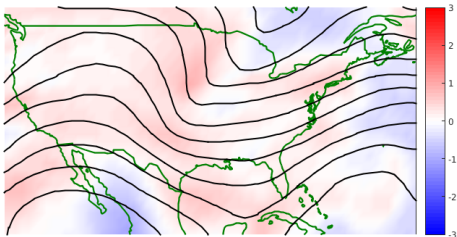
Winter storm evolution (U of Wisconsin)

2000 surprise snowstorm case.



ψ_{500}

F000



Outline

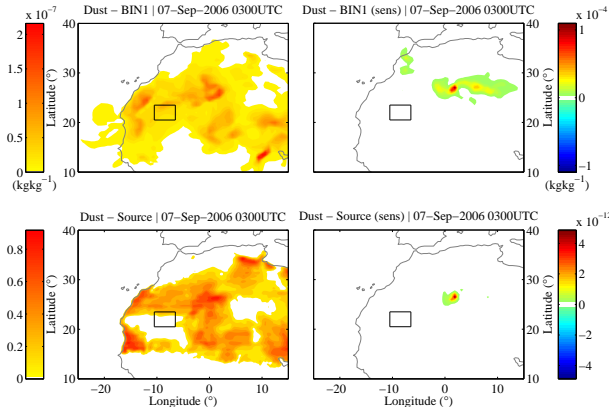
- Development of the GEOS-5 adjoint
- Applications
- Ongoing & future work

FV3 adjoint development

- Finish testing the non-hydrostatic version of the code.
- Generate adjoint of nesting, which requires different MPI communication.
- Develop system for maintaining the checkpoints through 4DVAR minimization.
- Increase efficiency.

GEOSctm Adjoint Model

GMAO have recently developed a chemistry transport version of GEOS-5. Having an adjoint of this mostly linear system will allow for long window estimation of sensitivity to emissions.



Singular vector development

We have recently 'revived' the singular vector system in GEOS-5 and development work is ongoing.

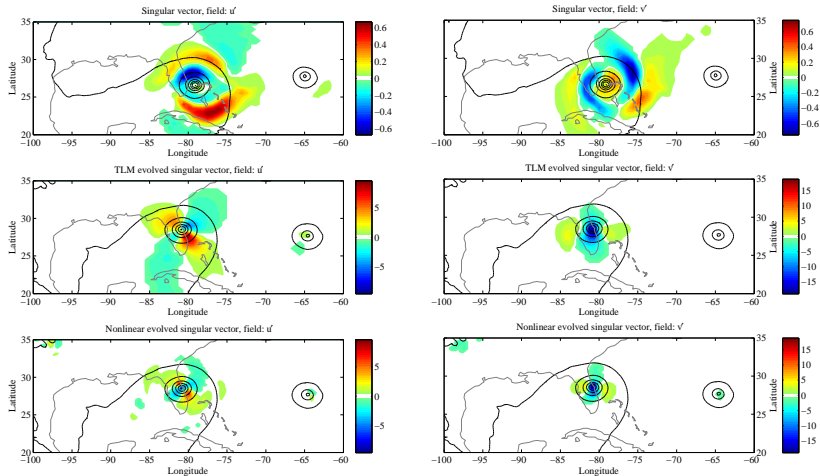
- Continue testing the singular vector system comparing to the literature
- Remove current restriction of horizontally global norm at initial time
- Add further norms useful for examining tropical cyclone predictability

Singular vector research

- Compare singular vector growth with ensemble spread used in 4DEnVar.
- Apply the singular vector tool to the SSW problem.
- Examine the low predictability hurricane Matthew case and the interaction with Nicole.

Singular vector research - hurricane Matthew

12 hour dry leading singular vector growth, total energy norm.



Hybrid 4DVAR

We have begun testing the latest FV3 based adjoint in a hybrid 4DVAR system and hope to start examining results soon.

Later we plan to make the adjoint of FV3 available to the community through the NGGPS and JEDI where testing of 4DVAR can continue and be compared to the current 4DEnVar techniques used at NCEP.

Efficiency gains to the adjoint will be key to its success in the 4DVAR framework. We are working with computer scientists to at GMAO to identify ways to improve performance.

Weak constraint?!

Conclusions

GMAO have developed sophisticated tangent linear and adjoint versions of the GEOS-5 global NWP model.

- Can be kept up to date with changes to FV3 using Tapenade approach.
- The system is scientifically sound but needs to be sped up.
- Lots of interesting research questions to address, especially with the nesting.